Marine transmissible cancer navigates urbanised waters, threatening to spillover

Maurine Hammel, Fanny Touchard, E a V Burioli, Laure Paradis, Frédérique Cerqueira, E. Chailler, I. Bernard, H. Cochet, A. Simon, Frédéric Thomas, et al.

Marine transmissible cancer navigates urbanised waters, threatening to spillover

Hammel M. [1,2], Touchard F. [1], Burioli E. A. V. [1,2], Paradis L.[1], Cerqueira F[1]., Chailler E.[1], Bernard I.[3], Cochet H.[4], Simon A.[1], Thomas F.[5], Destoumieux-Garzón D.[2], Charrière G. M.[2], Bierne N.[1]

[1] ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France

[2] IHPE, Univ Montpellier, CNRS, Ifremer, Univ Perpignan, Perpignan, France

[3] Eurêka Mer, Lézardrieux, France

[4] Cochet Environnement, 56550 Locoal

[5] CREEC/CANECEV (CREES), MIVEGEC, Unité Mixte de Recherches, IRD 224–CNRS 5290–Université de Montpellier, Montpellier, France

Abstract

Transmissible cancer, a unique form of microparasites that spreads through direct transmission of living cancer cells, is increasingly reported in marine bivalves. In this study, we sought to understand the ecology of the propagation of *Mytilus trossulus* Bivalve Transmissible Neoplasia 2 (MtrBTN2), a transmissible cancer affecting four *Mytilus* mussel species worldwide. We investigated the prevalence of MtrBTN2 in the mosaic hybrid zone of *M. edulis* and *M. galloprovincialis* along the French Atlantic coast, sampling contrasting natural and anthropogenic habitats. We observed a similar prevalence in both species, likely due to the proximity of the two species in this region. Our results showed that ports had higher prevalence of MtrBTN2, with a hotspot observed at a shuttle landing dock. No cancer was found in natural beds except for two sites around the hotspot, suggesting spillover. Ports may provide favourable conditions for the transmission of MtrBTN2, such as high mussel density, confined sheltered shores, or buffered temperatures. Ships may also spread the disease through biofouling, with maritime traffic being the best predictor of MtrBTN2 prevalence. Our results suggest ports may serve as epidemiological hubs, with maritime routes providing artificial gateways for MtrBTN2 propagation. This highlights the

29 importance of preventing biofouling on docks and ships' hulls to limit the spread of marine
30 pathogens.
31
32 **Keywords:** Transmissible cancer, Bivalve Transmissible Neoplasia, epidemiology,
33 biofouling, spillover, *Mytilus* mussels
34
35 **1. Introduction**
36
37 Transmissible cancers are malignant cell lineages that have acquired the ability to
38 infect new hosts through the transmission of living cancer cells. Eleven transmissible cancer
39 lineages have been described to date: one in dogs (canine transmissible venereal tumour,
40 CTVT [1, 2]), two in Tasmanian devils (devil facial tumour, DFT1 and DFT2 [3, 4]), and
41 eight in different marine bivalve species (bivalve transmissible neoplasia, BTNs [5–9]).
42 While direct contact is necessary for CTVT and DFTs transmission, via coitus for the former
43 and biting for the latter, the transmission of BTNs between sessile shellfish is assumed to
44 occur through the water column [10, 11]. *Mytilus* spp. mussels are affected by two BTN
45 lineages known as MtrBTN1 and MtrBTN2 that originate from two different *M. trossulus*
46 founder hosts. MtrBTN2 is distributed worldwide and has crossed the species barrier several
47 times to infect four *Mytilus* species: *M. trossulus* in East Asia [12, 13], *M. chilensis* in South
48 America [9], and *M. edulis* and *M. galloprovincialis* in Europe [9, 14]. The oldest MtrBTN2
49 sample is 14 years old. Genetic data suggest that it could be at least one or two orders of
50 magnitude more ancient, although dating is difficult due to an apparent acceleration of the
51 mitochondrial clock [12, 14].
52
53 The global distribution of the MtrBTN2 lineage remains enigmatic [9]. The other known
54 transmissible cancer with such a worldwide distribution is CTVT in dogs. CTVT emerged
55 4,000 to 8,000 years ago, probably in Eastern Asia, and has expanded rapidly worldwide over
56 the past 500 years, probably facilitated by the development of maritime transportation [15].
57 Transmission of MtrBTN2 within mussel populations occurs presumably through filter-
58 feeding (discussed in [10, 16]). Cancer cells can survive for several days in seawater [10],
59 which greatly increases their chances of infecting new hosts. However, it is unlikely that
60 transport of cancer cells by marine currents alone explains such a global distribution. Indeed,
61 just as gamete dispersal effectiveness is limited to a few metres in free-spawning
62 invertebrates [17], the dilution of cancer cells in the marine environment may strongly limit

the success of transmission over long distances (even locally in a mussel bed). Shipping traffic has been proposed as the most likely explanation for the global distribution of MtrBTN2 -i.e., the transport of disease-carrying mussels on ship hulls [9, 12]. To our knowledge, these claims have remained speculative and no formal evidence of the effect of shipping traffic and port habitat on the epidemiology of BTN has been provided to date.

European mussels are affected by MtrBTN2, with a much higher prevalence in *M. edulis* (of the order of 1/100) than in *M. galloprovincialis* (of the order of 1/1000) [14]. These two host species are parapatric in Europe and can coexist in contact zones where hybridisation is taking place [18–20]. Associations between genetic backgrounds and environmental variables indicate that the environment partly influences the structure and maintenance of the hybrid zones: in the study area, *M. edulis* genotypes are more frequent in sheltered habitats under freshwater influence, while *M. galloprovincialis* genotypes are more frequent in habitats exposed to wave action and ocean salinities [21, 22]. In this context, the distribution and propagation of MtrBTN2 is likely to be influenced by host genetic background and population composition, as well as by other environmental factors such as pollution, population density and abiotic environmental variables (e.g., salinity, temperature, pH). However, little is known about the ecology of MtrBTN2 in relation to its host and the possible environmental factors influencing the epidemiology of this transmissible cancer.

In the present study, we investigated the occurrence of MtrBTN2 in a survey area in Southern Brittany (France) which has contrasting natural and anthropogenic habitats and a mosaic distribution of *M. edulis* and *M. galloprovincialis*. We sampled 40 natural populations as well as 9 mussel farms, 7 floating buoys, and 20 ports. To detect MtrBTN2-infected mussels, we developed and performed a genetic screening method using pooled samples analysed by digital PCR (ddPCR) followed by demultiplexed real-time PCR. We amplified one nuclear and two mitochondrial genes using newly designed primers targeting specific variants. We found a low prevalence of MtrBTN2 (23/1516), with only 9 of the 76 sites of the study area affected, including 7 ports (the commercial port, 4 of the 16 marinas and 2 of the 3 landing docks). In contrast to our previous report [14], *M. galloprovincialis* were equally affected, possibly because they coexist closely with *M. edulis* in this region. Interestingly, one site had a high number of MtrBTN2 cases (8/20) and appeared to be a possible hotspot. Given that this site is a landing pontoon for maritime shuttles with frequent shipping traffic to other ports of the area, it could be a source of propagation. A few cancers were found in natural

beds near the hotspot, suggesting spillover, although none were found in sites outside other more enclosed ports where MtrBTN2 was detected. In addition, mussel farms and buoy samples were all MtrBTN2-free. These observations suggest that ports could be favourable habitats for MtrBTN2 transmission and/or that maritime transport could play a role as a vector for the spread of the disease, highlighting the role of ports as epidemiological hubs for MtrBTN2 propagation.

2. Materials and methods

2.1. Sampling

We collected 1516 *Mytilus* spp. mussels from 76 sites of the French Atlantic coast, from the Bay of Quiberon to Pornic, between January and March 2020. We collected hemolymph from adductor muscle (with a 1 ml syringe, 26G needle) and a piece of mantle; both samples were fixed with 96% ethanol as described in Hammel et al. [14]. DNA extraction was performed with the NucleoMag 96 Tissue kit (Macherey-Nagel) using a Kingfisher Flex extraction robot (serial number 711-920, ThermoFisher Scientific). We followed the kit protocol with a modified lysis time of 1 hour at 56°C and modified the volumes of some reagents: beads were diluted twice in water, 200 µl of MB3 and MB4, 300 µl of MB5 and 100 µl of MB6. DNA concentration (ng/µl) was quantified using the Nanodrop 8000 spectrophotometer (ThermoFisher Scientific). DNA from hemolymph was used for MtrBTN2 detection and DNA from mantle was used for mussel genotyping (Figure S1).

2.2. MtrBTN2 detection

As the prevalence of MtrBTN2 is expected to be low [14, 23] and the number of samples is high (n = 1516), we chose to pre-screen by pooling 12 samples (pooled screening step; Figure S1) and then demultiplexing positive pools to specifically target cancerous samples (simplex screening step; Figure S1). Prior to pooling, DNA concentrations were adjusted to 10 ng/µl to obtain an unbiased representation of each sample in the pool. This allowed us to considerably reduce the cost and time of detection, but we acknowledge that this protocol could result in some very early stages of the disease being missed. To partly circumvent this problem, MtrBTN2 detection in pools was performed by digital PCR (ddPCR) targeting one nuclear (Elongation Factor, EF) and one mitochondrial (Control region, described in Yonetmitsu et al. [9]) marker (Table S1). ddPCR is a sensitive method, based on the Taqman method which

requires two primers and one probe, and which directly estimates the copy number of the targeted sequence. In addition, the use of a mitochondrial marker should provide a more sensitive detection as the mitochondrial genome has more copies than the nuclear genome. ddPCR primers design and analysis was subcontracted to the company IAGE using the QIAcuity™ system. Positive pools were then demultiplexed (simplex screening step; Figure S1) and MtrBTN2 detection was performed by real-time PCR targeting one nuclear (Elongation Factor, EF1α i3) and one mitochondrial (cytochrome c oxidase I, mtCOI-sub) marker (Table S1).We used sequences of *M. edulis*, *M. galloprovincialis*, *M. trossulus*, and MtrBTN2 available from the National Center for Biotechnology Information (NCBI) to design the primer pairs (EF1α i3 and mtCOI-sub). Amplification of both markers was performed using a three-step cycling protocol (95°C for 10 s, 58°C for 20 s, 72°C for 25 s) for 40 cycles. We carried out real-time PCRs using the sensiFAST SyBR No-ROX Kit (Bioline) and the LightCycler 480 Real-Time PCR System. We also confirmed the positive samples diagnosed by real-time PCR using ddPCR (simplex screening step; Figure S1). The specificity of both method and markers was tested using various control samples of *M. edulis*, *M. galloprovincialis*, *M. trossulus*, and MtrBTN2 samples of early, moderate, or late stages (based on cytological observations, following Burioli et al. [23], Tableau S2). To test the sensitivity, we used positive controls with MtrBTN2 DNA diluted 10-fold (x10) and 100-fold (x100) in *M. edulis* DNA.

Real-time PCR results were analysed using the LightCycler480 software. Two parameters were extracted: threshold cycle (CT), obtained by Absolute Quantification analysis using the Second Derivative Maximum method, and melting temperature (Tm), obtained by Tm calling analysis. Positive samples were defined as those with a CT of less than 35 cycles and a Tm between 77.46 to 78.71 (based on Tm values of the positive control samples; see result, Figure S2). ddPCR results were analysed with the QuantaSoft™ Analysis Pro software, which provides the copy number for each sample. Samples were considered positive when there were more than two positive droplets.

2.3. Mussel genotyping

We used 10 biallelic SNPs known to discriminate between *M. edulis* and *M. galloprovincialis* mussel species (Table S3). These markers were identified as being ancestry-informative (fixed for alternative alleles in *M. edulis* and *M. galloprovincialis*) in the Fraïsse et al. dataset [24] and were subsequently confirmed as near diagnostic by analysis of larger datasets [20,

25]. Genotyping was performed using the Kompetitive Allele Specific PCR (KASP) method [26, 27]. As we wanted to know the mean ancestry of each sample (G ancestry) and reduce the cost of genotyping, we developed a multiSNPs marker by multiplexing the 10 SNPs. Rapidly, 1µL of assay mix (KASP-TF V4.0 2X Master Mix, 1X target concentration; primers, 1µM target concentration; HyClone™ HyPure water) and 0.5 µL of DNA at 10 ng/µL were mixed in qPCR 384-well plates using Labcyte Echo525. KASP analysis was performed on the LightCycler 480 Instrument (ROCHE) with the following thermal cycling conditions: initialisation 15 min at 95°C, first amplification 20 s at 94°C and 1 min at 61°C to 55°C with steps of 0.6°C for 10 cycles, second amplification 20 s 94°C and 1 min at 55°C for 29 cycles, read 1 min at 37°C and 1 s at 37°C.

KASP results are a combination of two fluorescence values, one for allele X, and another for allele Y. Here, we oriented all SNPs so that allele X corresponds to *M. edulis* and allele Y to *M. galloprovincialis*. Following Cuenca et al. [28], we transformed the data to obtain a single measure of the relative fluorescence of the two alleles, using the following formula:

$$y_{i,j} = Y_{i,j}/(X_{i,j} + Y_{i,j}),$$

where $X_{i,j}$ and $Y_{i,j}$ are the fluorescence value of allele X and allele Y respectively for i^{th} SNP of the j^{th} sample. To scale $y_{i,j}$ value from 0 when *M. edulis* allele fluorescence dominates at the 10 SNPs to 1 when *M. galloprovincialis* allele fluorescence dominates, we used the following formula:

$$y'_{i,j} = (y_{i,j} - \min(y_i))/(\max(y_i) - \min(y_i)),$$

where $y_{i,j}$ is the relative fluorescence value for the i^{th} SNP of the j^{th} sample. In the results section and figures, $y'_{i,j}$ corresponds to "G ancestry".

2.4. Environmental data

The 76 sampling sites were selected to present contrasting environments and genetic backgrounds (*M. edulis*, *M. galloprovincialis* and hybrids). The description of environmental data considered in this study is reported in Table S4 and all site information is available in Table S5. During sampling, we collected several habitat discrete descriptive variables: type of site (port, natural bed, farm, buoy), population density (1: isolated, 2: clustered, 3: beds), and wave action (1: exposed, 2: sheltered). We used the E.U. Copernicus Marine Service Information (doi: 10.48670/moi-00027) to access temperature (°C), salinity (PSU 1e-3) and current velocity (m.s-1) data for each site (monthly recorded from 2021, resolution of 0.028°x0.028°). We also used EMODnet Human Activities (source: EMSA Route Density

Map) to access the maritime traffic density of fishing, passenger, and recreational boats in the studied area (seasonal record from 2018, 2019 and 2020, resolution 1x1 km). In the analysis we used the annual average of salinity, current and boat density (maritime traffic). For temperature, we used two measures: the average of winter and autumn temperatures (TempW) and the average of summer and spring temperatures (TempS) as we identified opposite effects between winter and summer temperatures that would be misinterpreted if only the annual average temperature was used (Figure S3).

2.5. Statistical analysis: relationship between MtrBTN2 prevalence, host genotypes and environmental variables

All statistical analyses were performed in R *v4.1.3* [29]. Fisher's exact tests were used for the analysis of contingency tables of cancer prevalence (*fisher.test()* function of the *stats* package *v4.1.3* [29]). A principal component analysis (PCA) was used to explore the relationship between the quantitative environmental variables (*PCA()* function of the *FactoMineR* package *v2.6* [30]; missing values were estimated using *imputePCA()* of the *missMDA* package *v1.18* [31]). The association of MtrBTN2 prevalence with each environmental variable (Table S4) was tested using a univariate Poisson regression (*glm()* function of *stats* package *v4.1.3* [29]). Significant variables (p-value < 0.05) were selected and included in a multivariate Poisson regression. We also used redundancy analysis (RDA) to assess the overall relationship between cancer diagnosis and environmental variables. Real-time PCR and ddPCR results for both markers were synthesised to define positive samples: negative samples had 0 and cancerous samples had CT or copy number values (1516 x 4 matrix). All environmental data described above, including site types, were used as explanatory variables (*rda()* function of the *vegan* package *v2.6-4* [32]). Forward selection using the *ordiR2step()* *vegan* function was then used to determine the significance of the variables.

3. Results

3.1. Validation of the screening method

In order to detect MtrBTN2 in thousands of individuals, we developed dedicated ddPCR and qPCR screening tools.

The design of real-time PCR and ddPCR primers and probes allowed us to amplify nuclear and mitochondrial loci specific to *M. trossulus* mussels and MtrBTN2 (Figure S4, TableS6). However, for the mtCOI-sub locus, the melting temperature (Tm) range of MtrBTN2 was

different from *M. trossulus* controls and allowed us to specifically detect MtrBTN2 based on Tm values (Figure S2).

With ddPCR, all MtrBTN2 control samples were detected, regardless of marker or dilution, except for one MtrBTN2 sample diluted at x100 which was negative for the EF locus (Figure S4). Copy numbers appeared higher for the mtCR marker, showing a higher sensitivity for this mitochondrial marker. The same pattern was observed for EF1α-i3 and mtCOI-sub in real-time PCR, with a higher sensitivity of the mitochondrial marker based on the lower CT values observed (Figure S4). However, two MtrBTN2 samples diluted at x100 were negative for mtCOI-sub and positive for EF1α-i3. Even though the mitochondrial marker appears to be more sensitive in both methods, we chose to screen for MtrBTN2 using both markers in order to have comparisons and additional validation.

3.2. MtrBTN2 screening reveals a low prevalence

We amplified nuclear and mitochondrial markers using ddPCR (EF, mtCR) and real-time PCR (EF1α-i3, mtCOI-sub) to detect the presence of MtrBTN2 in the 1516 mussels sampled from the 76 sites of the study area (Figure S1). All results of the pooled and simplex screening are presented in TableS7.

The pooled screening step using ddPCR revealed 25/127 positive pools: 9/127 with EF, 14/127 with mtCR, and 2/127 were positive for both markers (Figure S4). As the 14 mtCR-positive pools showed low copy numbers (median = 7.59), we assume that EF did not amplify because of a detection threshold. For pools positive for both markers, the EF copy numbers appear relatively low compared to mtCR copy numbers. This is due to the high mitochondrial fluorescence masking the nuclear fluorescence and underestimating EF copy numbers. To eliminate this bias in the simplex ddPCR step (Figure S1), we amplified the two markers separately. In order to maximise cancer detection in the simplex screening step, we chose to keep the pools positive for at least one gene, corresponding to 25 positive pools of 12 samples (n = 300).

We performed the simplex screening step using real-time PCR and then confirmed positive samples with ddPCR (Figure S1). In total, 44/300 samples were positive for at least one real-time PCR marker: 16 with EF1α-i3 only, 23 with mtCOI-sub only, and 5 with both markers (Figure S4, Table S8). Pooled and simplex screening results are mostly consistent, as 20 of the 25 positive pools had at least one individual detected by real-time PCR amplification (Table S9). The 5 pools without positive samples had low ddPCR copy numbers (less than 3.37 for EF and less than 6.65 for mtCR). These false positives were to be expected as we

intentionally applied low threshold values in the pooled screening step. Two samples from mtCR-positive pools were found to be positive for both markers at the simplex level, which was expected as the mitochondrial marker is likely to be more sensitive than the nuclear marker. On the contrary, we did not expect to observe samples positive for EF1α-i3/EF and negative for mtCOI-sub/mtCR. However, all but one of the individuals from the EF-positive pools were positive only for the EF1α-i3 marker (Table S9). The absence of mitochondrial genes in these samples does not seem to be related to a detection threshold issue, especially as ddPCR and real-time PCR target two different genes (mtCR and mtCOI-sub). We suspect that some host alleles that were present at low frequencies in the 'dock mussels' populations (the hybrid lineage found in the port of Saint-Nazaire [25]) may sometimes be amplified with EF1α-i3. This could be explained by shared polymorphism between host species rather than the presence of MtrBTN2. Therefore, the 16 samples positive only for EF1α-i3 were conservatively considered as false positives. For the remaining 28 samples, we performed ddPCR with both markers to confirm the real-time mitochondrial PCR diagnosis (Figure S1, TableS8). For 23/28 samples, mtCR ddPCR confirmed the mtCOI-sub real-time PCR diagnosis. For 10/23 samples positive only for mtCOI-sub, ddPCR revealed few EF copies. Finally, we considered all samples positive for mitochondrial markers with both methods to be cancerous samples. We found 23/1516 positive samples from 9/76 sites in which the number of cases varied from 1 to 8 (Figure 1). The highest proportion of cancerous mussels was found at site P20 (Croisic Pontoon), and the next highest in the vicinity of this site (P19, P21, and P23; Figure 1).

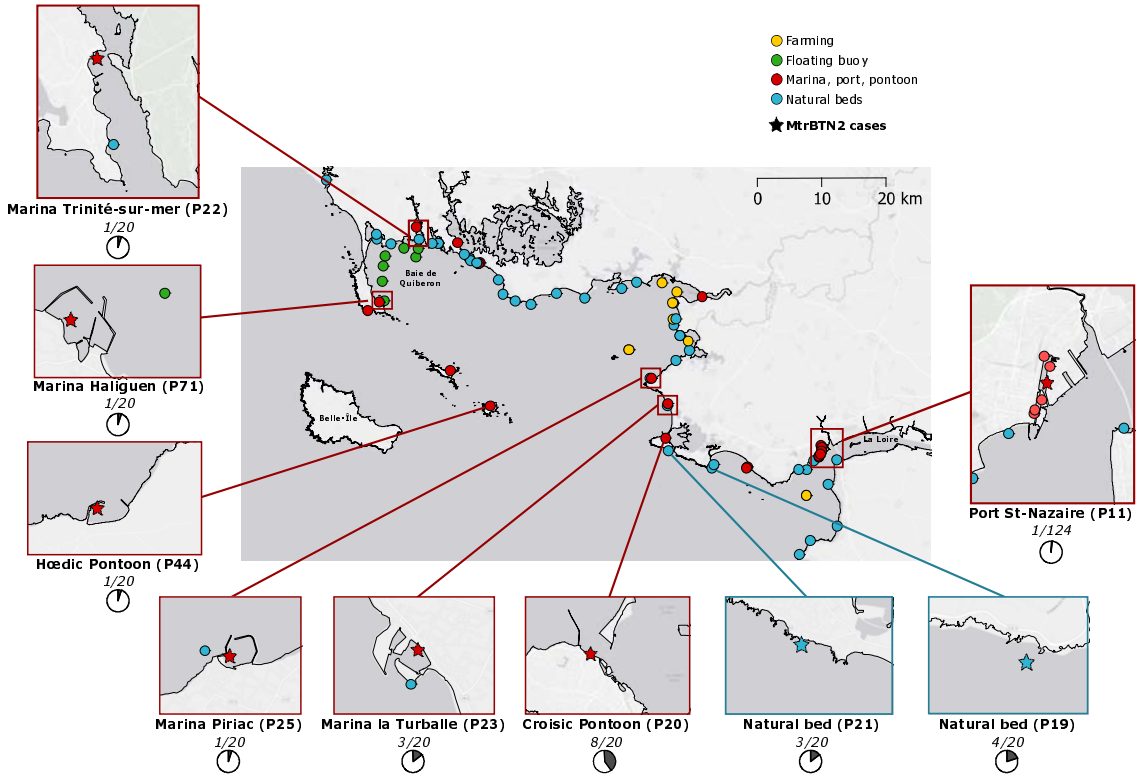


Figure 1: Location of sampling sites. Zoomed-in sites correspond to sites where mussels are affected by MtrBTN2. Pie graphs and numbers in italics correspond to the proportion of cancerous samples in each site. Coloured points represent the type of site, as indicated in the figure's legend. Map origin: Ersi gray (light).

3.3. M. edulis and M. galloprovincialis are equally affected by MtrBTN2

To determine whether host genetic background or population composition influences cancer prevalence and distribution, mussels were genotyped using 10 SNPs known to be diagnostic between *M. edulis* and *M. galloprovincialis* (the two species coexisting in the studied area).

To reduce the cost of genotyping, we developed a multiplex assay, mixing 10 SNPs that directly give the average ancestry of the sample (G ancestry, Figure S1). This approach was validated by comparing the multiplex fluorescence and average simplex fluorescence of SNPs from a set of samples (Figure S5a). We genotyped each individual from the 9 sites affected by MtrBTN2 (9*20 = 180 mussels) and estimated the average ancestry of each site by genotyping pools of individuals (76 pools of 20 mussels). The good positive correlation between the average individual fluorescence of a site and the fluorescence of the corresponding pool validated our approach (Figure S5b). Mussels with G ancestry values below 0.3 were assigned to *M. edulis*, those with values above 0.59 to *M. galloprovincialis*, and those in between were assigned as hybrids (Figure 2, FigureS6). Of the mussels affected

by MtrBTN2, 15/23 were assigned to *M. edulis*, 7/23 to *M. galloprovincialis* and 1/23 was a hybrid (Table S10). Considering the sites affected by MtrBTN2, MtrBTN2 prevalence did not appear to be significantly different between species with 15/105 (14%) in *M. edulis*, 7/60 (11%) in *M. galloprovincialis* and 1/15 (6%) in hybrids (Fisher exact test, p-value = 0.83). Most *M. edulis* mussels affected by MtrBTN2 were found in populations with a majority of *M. edulis* (P44, P23, P20), only two were found in populations with some hybrids and *M. galloprovincialis* mussels (P25, P71) and only one in a population with a majority of *M. galloprovincialis* (P11, Figure 2, Figure S6). Interestingly, all *M. galloprovincialis* mussels affected by MtrBTN2 were found in *M. galloprovincialis* populations (P19, P21, Figure 2, FigureS6), which are close to the site with the highest prevalence (Croisic pontoon, Figure 1). Finally, the hybrid individual affected by MtrBTN2 was found in a mixed population (P22, Figure 2, FigureS6).

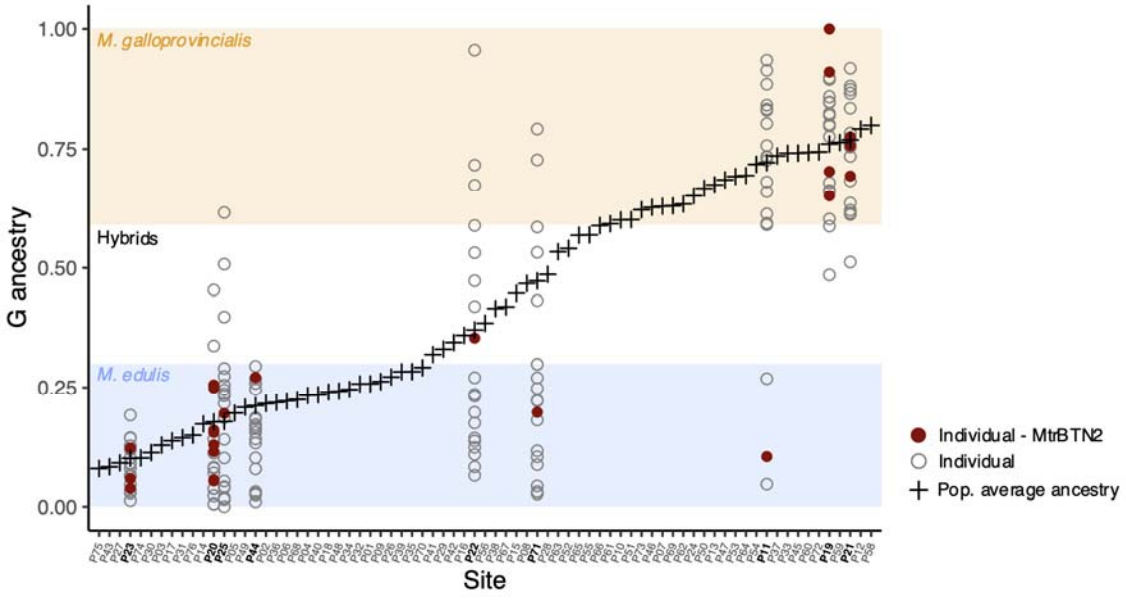


Figure 2: Individual and population ancestry. G ancestry of healthy (empty grey circle) and MtrBTN2 affected (dark red) hosts is shown for the 9 sites with cancer (bold labels). Population average ancestry for all sampled sites is represented by black crosses. Blue, white and orange rectangles correspond to G ancestry range of *M. edulis*, hybrids and *M. galloprovincialis* respectively.

3.4. MtrBTN2 infected mussels are mostly found in ports with high density of maritime traffic

After the MtrBTN2 detection step, which revealed 23/1516 individuals infected with this transmissible cancer, we investigated associations with environmental factors. Cancerous

samples were found in 7 ports (16/23 infected mussels) and 2 natural beds (7/23) near the Croisic pontoon (P20, Figure 1). MtrBTN2 was not detected in the 180 individuals from the 9 mussel farms or in the 140 individuals from the 7 floating buoys. Cancerous samples were significantly more frequent in ports than in other types of sites (Fisher exact test, p-value = 0.0001, Table S11).

The PCA analysis reveals an association between environmental variables and the types of sites rather than the presence of cancer itself (Figure 3). Indeed, ports, floating buoys and farms are well separated from natural beds on the first factorial plan while similarly found in sheltered habitats with high population density. Interestingly, maritime traffic (i.e., annual mean of passenger, fishing, and recreational boat density) tends to explain a direction of variation towards human-altered habitats, but also towards sites affected by MtrBTN2. MtrBTN2 seems to be found in sites with cooler mean temperatures in summer and warmer mean temperatures in winter and high mean annual salinity and current velocity, environmental variables that characterise ports. Water temperature tends to be buffered in ports, being warmer in winter and cooler in summer (Figure S3). However, these variables are also difficult to interpret as they also explain well the inertia of axis 1 of the PCA which differentiates the sites located near estuaries (La Loire or La Vilaine) from the others. To better identify the environmental variables that best explain MtrBTN2 prevalence despite being correlated, we used an ordination method (RDA) and a generalized linear model (GLM).

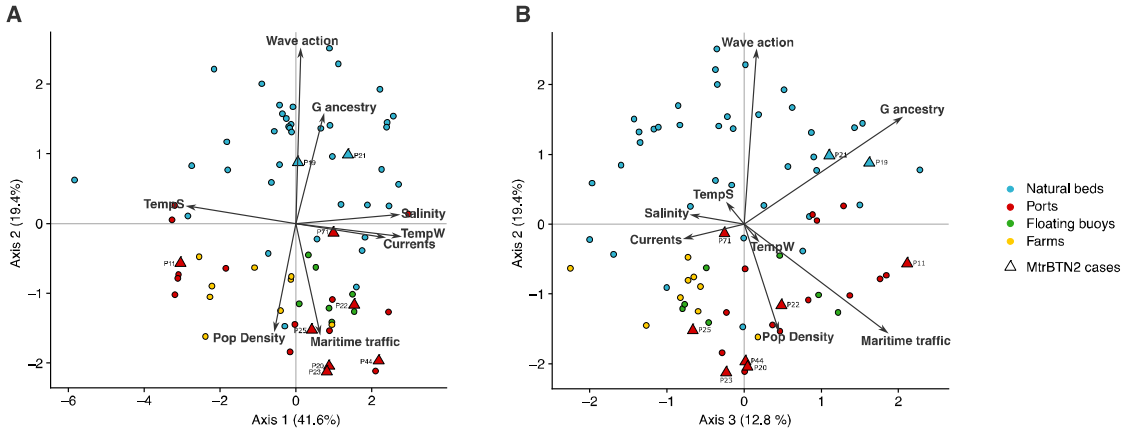


Figure 3: PCA on environmental variables. (A) axis 1 versus axis 2, (B) axis 2 versus axis 3 of the PCA analysis.

356 The RDA analysis explores the statistical association between the presence of cancer and
357 environmental variables, including types of sites. Of the 9 environmental variables included
358 in our study (Figure S7), 4 were found to be significant in explaining the presence of cancer
359 after bootstrapping: maritime traffic (p-value = 0.002, R^2adj = 0.013), population density (p-
360 value = 0.012, R^2adj = 0.015), site type (p-value = 0.02, R^2adj = 0.020) and tempS (spring-
361 summer temperature; p-value = 0.02, R^2adj = 0.022). It is worth noting that, according to the
362 PCA observation, tempS may explain the influence of estuary water and buffered temperature
363 in ports rather than the presence of cancer itself.
364 The multivariate Poisson GLM on cancer prevalence was conducted on the four numerical
365 variables that significantly explained the prevalence of cancer in the univariate models
366 (maritime traffic, population density, spring-summer temperature (tempS) and currents, Table
367 S12). Categorical data on site type were not included in this analysis. At the multivariate
368 level, the prevalence of cancer was significantly explained by maritime traffic (p-value =
369 0.002) and population density (p-value = 0.026, Table S12), in agreement with the RDA.
370 The univariate correlations of the four variables that best explain MtrBTN2 prevalence in the
371 ordination and GLM analyses, namely maritime traffic and population density, as well as site
372 type and tempS, are presented in Figure 4.
373

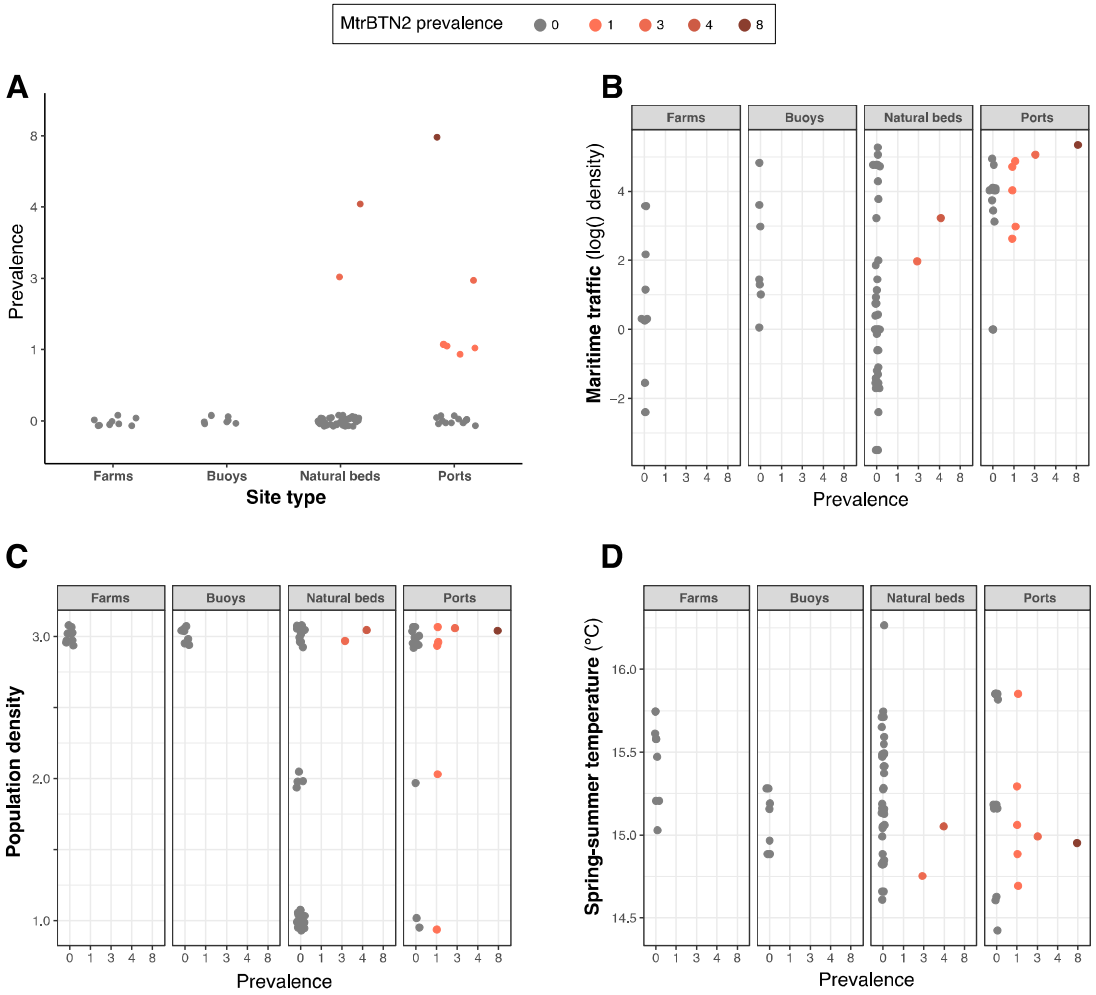


Figure 4: Univariate correlations of environmental variables that best explain the prevalence of MtrBTB2: site type (A), maritime traffic (B), population density (C) and spring-summer temperature (D). Coloured points correspond to the prevalence of MtrBTN2 as indicated in the legend.

4. Discussion

Our results show that MtrBTN2 is present at low prevalence along the French Atlantic coast (23/1516 = 1.5%). Strikingly, it is mostly found in ports and is equally distributed in *M. edulis* and *M. galloprovincialis*. Our analysis of environmental factors influencing the epidemiology of MtrBTN2 shows that ports are likely epidemiological hubs for MtrBTN2 propagation.

Most MtrBTN2-infected mussels were found in ports and a hotspot was found around the Croisic pontoon. Given the low prevalence of MtrBTN2 previously described [14, 23] and

confirmed in this study, we acknowledge that our sample sizes per site may be too small (n = 20) to assess MtrBTN2 prevalence with confidence at a site level. However, our careful overall sampling design, capturing replications of habitat features, has allowed us to test for correlations between environmental factors and uncover higher prevalence of MtrBTN2 in ports.

The highest prevalence of MtrBTN2 in ports could be explained by two non-exclusive hypotheses. Firstly, ports could offer a favourable environment as it is conducive to disease development and propagation. Indeed, ports are often polluted, confined and permanently immersed habitats (without tidal constraints) with high density of mussels. Please consider that our objective here is not to explain the emergence of a transmissible cancer, but its maintenance and spread. The emergence probably took place long before the industrial period and the mutagenic effect of pollutants does not have to be put forward to explain our results, but rather their effect on host immunity. Secondly, there is increased connectivity between ports due to maritime traffic and ports are connected to other sites (other ports or wild sites) through biofouling from vessels.

Pollution is a major environmental stressor that could affect host physiology and increase susceptibility to parasites, particularly by altering the immune response (reviewed in [33]). As filter-feeders, species of bivalve molluscs are particularly concerned [34–37]. However, in our study, most of the affected individuals (15/23) were found in three open sea sites known to be healthy and protected ecosystems with abundant resources (Croisic Pontoon, Natural beds 19 and 21 part of Natura 2000 area; Figure 1), suggesting that pollution alone cannot explain the higher prevalence of MtrBTN2 cancers in ports. Conversely, it can be hypothesised that a healthy, less stressful environment may increase the host permissiveness and allow MtrBTN2 (and possibly other BTNs) to persist at a higher detectable prevalence. Host nutrition can affect disease outcome by driving host immunity and parasite resource availability [38–40]. In invertebrates, increased host nutrition can lead to increased parasite load [41] and improved immunity against parasites [39, 40]. This balance could apply to mussel-MtrBTN2 interaction, and a healthy environment with abundant resources could increase host carrying capacity and/or slow down disease progression. A polluted environment could also affect MtrBTN2 cells survival in seawater and reduce their chances of infecting a new host. However, the impact of pollution on host and cancer cells needs to be

423 further investigated to clearly establish its influence on the transmission and persistence of
424 such transmissible cancers.
425
426 Ports are often confined areas with long particle residence times which could increase the
427 contact time between parasites and hosts, thereby increasing the probability of transmission
428 and favouring the persistence of a passively dispersing parasite. However, the major hotspot
429 on Croisic Pontoon, which is affected by strong tidal currents, again provides evidence
430 against this hypothesis given that it is not a confined area.
431
432 Host population density is an important epidemiological parameter, as high density is known
433 to favour parasite persistence in host populations [42–44]. Our results suggest that highest
434 MtrBTN2 prevalence rates are found in sites with high population density. However, we
435 found no evidence of MtrBTN2 cases in mussel farms despite high mussel population density
436 in these sites. Although population density could favour MtrBTN2 persistence in an affected
437 population, this parameter alone cannot explain the higher prevalence of MtrBTN2 in ports.
438
439 Connectivity through maritime traffic appears to be the main common feature between
440 affected ports. It is conceivable that certain hotspots exist at certain locations along the coast,
441 either due to environmental factors that have yet to be characterised, or simply due to
442 stochastic local outbreaks. If a port is affected by such a local outbreak, it will become a
443 source of contamination to other ports through biofouling of vessels, while natural beds will
444 probably remain isolated by density troughs that could halt the propagation (Figure 5). While
445 this does not disqualify the factors discussed above – e.g., pollution, population density and
446 confinement- as drivers of contamination within ports, our study suggests that maritime
447 routes between ports could be anthropogenic gateways to MtrBTN2 propagation.
448 Furthermore, despite a high density of maritime traffic around the floating buoys we sampled,
449 these sites remained free of MtrBTN2. This could indicate that in addition to cruising, ships
450 must dock for the spread of MtrBTN2 to be effective, as is the case in ports (Figure 5). This
451 result represents a significant advance in our understanding of MtrBTN2 epidemiology, both
452 by proposing ports as epidemiological hubs and by suggesting that the natural propagation of
453 this transmissible cancer could be naturally halted by density troughs in the distribution of
454 host mussel populations. *M. galloprovincialis* was thought to be more resistant to MtrBTN2
455 because of lower prevalence in this species [14], and patches of *M. galloprovincialis* could
456 also have acted as barriers to disease propagation between patches of *M. edulis*. However, our

results suggest that *M. galloprovincialis* population patches close to *M. edulis* populations may also show enhanced prevalence. This does not contradict the idea that *M. galloprovincialis* population patches could impose a slightly higher resistance to MtrBTN2 propagation. Under the hypothesis that natural propagation is hindered by density troughs between population patches, or by population patches of a more resistant species, ports and ship traffic would create anthropogenic gateways that favour connections between patches of host populations (Figure 5). This hypothesis would also suggest that BTNs spread naturally and progressively from one host to another by close contact rather than drifting with the currents, making enhanced R0 by anthropogenic dispersal a threat to host populations.

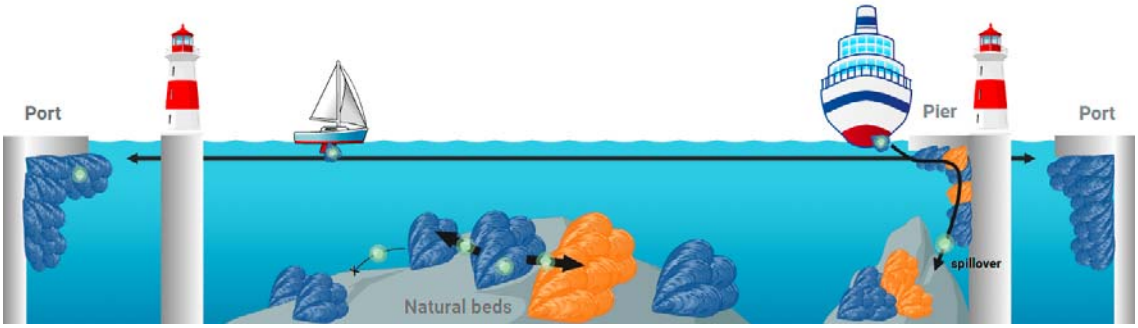


Figure 5: Diagram illustrating the discussion on the propagation of MtrBTN2 across *Mytilus* sp. populations. Arrows widths correspond to the probability of MtrBTN2 propagation. Blue and orange mussels correspond to *M. edulis* and *M. galloprovincialis*, respectively. Diagram created with BioRender.

A hotspot of MtrBTN2 prevalence was found in the area around the Croisic pontoon. This specific site could represent a source for other ports and surrounding natural beds. Indeed, this pontoon is frequently used by maritime shuttles, as well as by recreational and fishing boats, and is connected by maritime traffic to many ports in the studied area. In particular, this pontoon is the departure site of boats travelling to Hoedic (site P44 in Figure 1), an isolated island for which the presence of cancer is most likely explained by anthropogenic transport given the distance from the mainland. MtrBTN2 is also found in foreshores near Croisic. This reveals a possible spillover from ports to adjacent wild populations (Figure 5). Maritime traffic has been identified as a vector for non-indigenous species, including their expansion from commercial ports to natural populations via marinas [45–48]. It has also been documented that biofouling from ships can introduce parasites into unaffected areas and have secondary impacts on local populations, including causing significant mortality (see [49]). In our study, when sites outside standard enclosed ports (e.g., enclosed marinas) were sampled,

MtrBTN2 was never detected, suggesting that contagion remains confined to the port area. This suggests that spillover may depend on the degree of confinement of a port.

5. Conclusion

Our study reveals that the prevalence of MtrBTN2 along the South Britany French coast is low, but that the most affected sites are ports (marinas, commercial ports, and docks). This result suggests that ports are epidemiological hubs and maritime routes anthropogenic gateways for the propagation of MtrBTN2. Detection of MtrBTN2 presence on boat biofouling would provide more direct evidence of its role as vector. It would also be interesting to investigate whether prevalence hotspots are more commonly located in areas with particular environmental conditions, or whether they are stochastic local outbreaks. This could be done by detecting new hotspots or by sampling Croisic pontoon again in the future. The spreading of (micro)parasites and pathogens through vessel biofouling is a legitimate concern, and results such as ours -i.e., on a smaller scale than translocations identified in the marine bioinvasion literature or by Yonemitsu et al. [9]- highlight the critical need for politicy regulation to limit the effects of biofouling, both on ship hulls and port docks, in addition to ballast water control.

Data accessibility

All data used in this study are available in the supplementary material.

Author's contribution

M.H., Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Visualization, Writing – original draft, Writing – review and editing. F.To., Investigation, Data curation, Writing – review and editing. E.A.V.B., Formal analysis, Investigation, Writing – review and editing. L.P., Data curation, Investigation, Writing – review and editing. F.C., Investigation, Writing – review and editing. E.G., Investigation, Writing – review and editing. I.B., Resources, Writing – review and editing. H.C., Resources, Writing – review and editing. A.S., Formal analysis, Writing – review and editing. F.Th., Conceptualization, Funding acquisition, Writing – review and editing. D.D-G., Conceptualization, Funding acquisition, Writing – review and editing. G.M.C., Conceptualization, Project administration, Funding acquisition, Writing – review and editing. N.B., Conceptualization, Funding acquisi-

tion, Formal analysis, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review and editing.

Conflict of interest declaration

We declare we have no competing interests.

Fundings

This work was supported by Montpellier Université d'Excellence (BLUECANCER project) and Agence Nationale de la Recherche (TRANSCAN project, ANR-18-CE35-0009). This study falls within the framework of the "Laboratoires d'Excellence (LABEX)" CEMEB (10-LABX-0004) and Tulip (ANR-10-LABX-41). FT is supported by the MAVA Foundation.

Acknowledgement

We are grateful to the IAGE company for carrying out the ddPCR assay of this study. We thank the GENSEQ and QPCR HAUT-DEBIT platforms for access to equipments and their expertise. We thank Cécile Perrin for her valuable reading of the manuscript.

REFERENCES

[1] Murgia C, Pritchard JK, Kim SY, et al. Clonal Origin and Evolution of a Transmissible Cancer. *Cell* 2006; 126: 477–487.

[2] Rebbeck CA, Thomas R, Breen M, et al. Origins and evolution of a transmissible cancer. *Evolution (N Y)* 2009; 63: 2340–2349.

[3] Pearse AM, Swift K. Allograft theory: Transmission of devil facial-tumour disease. *Nature* 2006; 439: 549.

[4] Pye RJ, Pemberton D, Tovar C, et al. A second transmissible cancer in Tasmanian devils. *Proceedings of the National Academy of Sciences* 2016; 113: 374–379.

[5] Garcia-Souto D, Bruzos AL, Diaz S, et al. Mitochondrial genome sequencing of marine leukaemias reveals cancer contagion between clam species in the Seas of Southern Europe. *Elife*; 11. Epub ahead of print 1 January 2022. DOI: 10.7554/ELIFE.66946.

[6] Metzger MJ, Reinisch C, Sherry J, et al. Horizontal transmission of clonal cancer cells causes leukemia in soft-shell clams. *Cell* 2015; 161: 255–263.

[7] Metzger MJ, Villalba A, Carballal MJ, et al. Widespread transmission of independent cancer lineages within multiple bivalve species. *Nature* 2016; 534: 705–709.

[8] Michnowska A, Hart SFM, Smolarz K, et al. Horizontal transmission of disseminated neoplasia in the widespread clam Macoma balthica from the Southern Baltic Sea. *Mol Ecol* 2022; 31: 3128–3136.

[9] Yonemitsu MA, Giersch RM, Polo-Prieto M, et al. A single clonal lineage of transmissible cancer identified in two marine mussel species in South America and Europe. *Elife*; 8. Epub ahead of print 1 November 2019. DOI: 10.7554/eLife.47788.

[10] Burioli EAV, Hammel M, Bierne N, et al. Traits of a mussel transmissible cancer are reminiscent of a parasitic life style. *Scientific Reports 2021 11:1* 2021; 11: 1–11.

[11] Giersch RM, Hart SFM, Reddy SG, et al. Survival and Detection of Bivalve Transmissible Neoplasia from the Soft-Shell Clam Mya arenaria (MarBTN) in Seawater. Epub ahead of print 2022. DOI: 10.3390/pathogens11030283.

[12] Skazina M, Odintsova N, Maiorova M, et al. First description of a widespread Mytilus trossulus-derived bivalve transmissible cancer lineage in M. trossulus itself. *Sci Rep*; 11. Epub ahead of print 2021. DOI: 10.1038/s41598-021-85098-5.

[13] Skazina M, Odintsova N, Maiorova M, et al. Two lineages of bivalve transmissible neoplasia affect the blue mussel Mytilus trossulus Gould in the subarctic Sea of Okhotsk. *Curr Zool* 2022; 00: 1–12.

[14] Hammel M, Simon A, Arbiol C, et al. Prevalence and polymorphism of a mussel transmissible cancer in Europe. *Mol Ecol* 2022; 31: 736–751.

[15] Baez-Ortega A, Gori K, Strakova A, et al. Somatic evolution and global expansion of an ancient transmissible cancer lineage. *Science (1979)*; 365. Epub ahead of print 2 August 2019. DOI: 10.1126/science.aau9923.

[16] Caza F, E B, FJ V, et al. Hemocytes released in seawater act as Trojan horses for spreading of bacterial infections in mussels. *Sci Rep* 2020; 10: 19696–19696.

[17] D.R. Levitan. Sperm Competition and Sexual Selection in External Fertilizers. In: Birkhead TR, MAP (ed) *Sperm competition and sexual selection*, pp. 175–217.

[18] Skibinski DOF, Beardmore JA, Cross TF. Aspects of the population genetics of Mytilus (Mytilidae; Mollusca) in the British Isles. *Biological Journal of the Linnean Society* 1983; 19: 137–183.

[19] Coustau C, Renaud F, Delay B. Genetic characterization of the hybridization between Mytilus edulis and M. galloprovincialis on the Atlantic coast of France. *Mar Biol* 1991; 111: 87–93.

[20] Simon A, Fraïsse C, el Ayari T, et al. How do species barriers decay? Concordance and local introgression in mosaic hybrid zones of mussels. *J Evol Biol* 2021; 34: 208–223.

[21] Bierne N, Bonhomme F, David P. Habitat preference and the marine-speciation paradox. *Proc R Soc Lond B Biol Sci* 2003; 270: 1399–1406.

[22] Bierne N, David P, Langlade A, et al. Can habitat specialisation maintain a mosaic hybrid zone in marine bivalves? *Mar Ecol Prog Ser* 2002; 245: 157–170.

[23] Burioli EAV, Trancart S, Simon A, et al. Implementation of various approaches to study the prevalence, incidence and progression of disseminated neoplasia in mussel stocks. *J Invertebr Pathol* 2019; 168: 107271.

[24] Fraïsse C, Belkhir K, Welch JJ, et al. Local interspecies introgression is the main cause of extreme levels of intraspecific differentiation in mussels. *Mol Ecol* 2016; 25: 269–286.

[25] Simon A, Arbiol C, Nielsen EE, et al. Replicated anthropogenic hybridisations reveal parallel patterns of admixture in marine mussels. *Evol Appl* 2020; 13: 575–599.

[26] Cuppen E. Genotyping by Allele-Specific Amplification (KASPar). *CSH Protoc* 2007; 2007: pdb.prot4841.

[27] Semagn K, Babu R, Hearne S, et al. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application in crop improvement. *Molecular Breeding* 2014; 33: 1–14.

[28] Cuenca J, Aleza P, Navarro L, et al. Assignment of SNP allelic configuration in polyploids using competitive allele-specific PCR: application to citrus triploid progeny. *Ann Bot* 2013; 111: 731–742.

[29] R Core Team. R: A language and environment for statistical computing, http://www.r-project.org/index.html (2020).
[30] Lê S, Josse J, Husson F. FactoMineR: An R Package for Multivariate Analysis. *J Stat Softw* 2008; 25: 1–18.
[31] Josse J, Husson F. missMDA: A Package for Handling Missing Values in Multivariate Data Analysis. *J Stat Softw* 2016; 70: 1–31.
[32] Oksanen J, Blanchet FG, Friendly M, et al. Package 'vegan' - Community Ecology Package Version 2.5-7.
[33] Destoumieux-Garzón D, Mavingui P, Boetsch G, et al. The one health concept: 10 years old and a long road ahead. *Front Vet Sci* 2018; 5: 14.
[34] Canesi L, Ciacci C, Bergami E, et al. Evidence for immunomodulation and apoptotic processes induced by cationic polystyrene nanoparticles in the hemocytes of the marine bivalve Mytilus. *Mar Environ Res* 2015; 111: 34–40.
[35] Galloway TS, Depledge MH. Immunotoxicity in Invertebrates: Measurement and Ecotoxicological Relevance. *Ecotoxicology 2001 10:1* 2001; 10: 5–23.
[36] Girón-Pérez MI. Relationships between innate immunity in bivalve molluscs and environmental pollution. *ISJ* 2010; 7: 149–156.
[37] Renault T. Immunotoxicological effects of environmental contaminants on marine bivalves. *Fish Shellfish Immunol* 2015; 46: 88–93.
[38] Chandra RK. Nutrition, immunity and infection: From basic knowledge of dietary manipulation of immune responses to practical application of ameliorating suffering and improving survival. *Proc Natl Acad Sci U S A* 1996; 93: 14304–14307.
[39] Tseng M, Myers JH. The Relationship between Parasite Fitness and Host Condition in an Insect - Virus System. *PLoS One* 2014; 9: e106401.
[40] Pike VL, Lythgoe KA, King KC. On the diverse and opposing effects of nutrition on pathogen virulence. *Proceedings of the Royal Society B*; 286. Epub ahead of print 10 July 2019. DOI: 10.1098/RSPB.2019.1220.
[41] Cressler CE, Nelson WA, Day T, et al. Disentangling the interaction among host resources, the immune system and pathogens. *Ecol Lett* 2014; 17: 284–293.
[42] Bommarito C, Thieltges DW, Pansch C, et al. Effects of first intermediate host density, host size and salinity on trematode infections in mussels of the south-western Baltic Sea. *Parasitology* 2021; 148: 486–494.
[43] Hopkins SR, Fleming-Davies AE, Belden LK, et al. Systematic review of modelling assumptions and empirical evidence: Does parasite transmission increase nonlinearly with host density? *Methods Ecol Evol* 2020; 11: 476–486.
[44] Thieltges DW. Habitat and transmission – effect of tidal level and upstream host density on metacercarial load in an intertidal bivalve. *Parasitology* 2007; 134: 599–605.
[45] Floerl O, Inglis GJ, Dey K, et al. The importance of transport hubs in stepping-stone invasions. *Journal of Applied Ecology* 2009; 46: 37–45.
[46] Murray CC, Pakhomov EA, Therriault TW. Recreational boating: a large unregulated vector transporting marine invasive species. *Divers Distrib* 2011; 17: 1161–1172.
[47] Ulman A, Ferrario J, Occhpinti-Ambrogi A, et al. A massive update of non-indigenous species records in Mediterranean marinas. *PeerJ* 2017; 5: e3954.
[48] Guzinski J, Ballenghien M, Daguin-Thiébaut C, et al. Population genomics of the introduced and cultivated Pacific kelp Undaria pinnatifida: Marinas—not farms—drive regional connectivity and establishment in natural rocky reefs. *Evol Appl* 2018; 11: 1582–1597.
[49] Georgiades E, Scianni C, Davidson I, et al. The Role of Vessel Biofouling in the Translocation of Marine Pathogens: Management Considerations and Challenges .

Frontiers in Marine Sciences; 8. Epub ahead of print 2021. DOI: 10.3389/fmars.2021.660125.

Supplementary figures

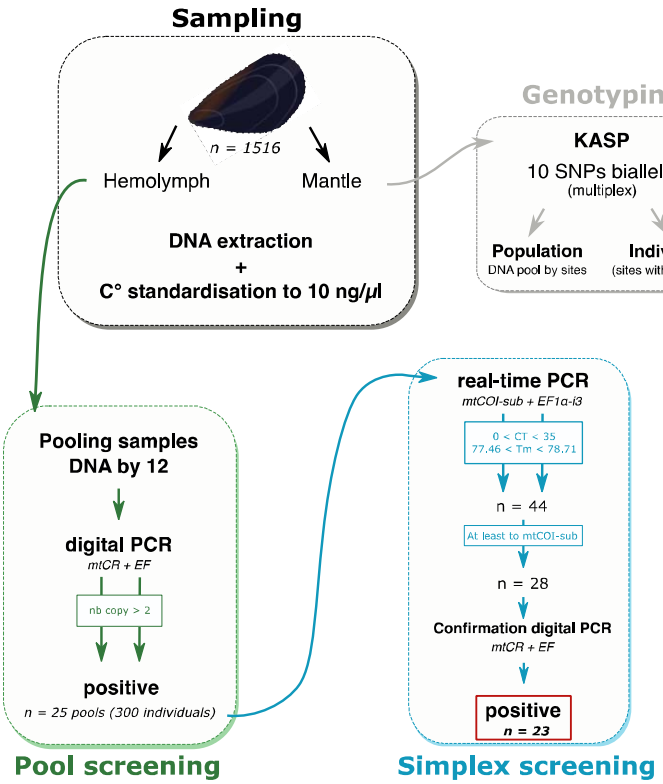
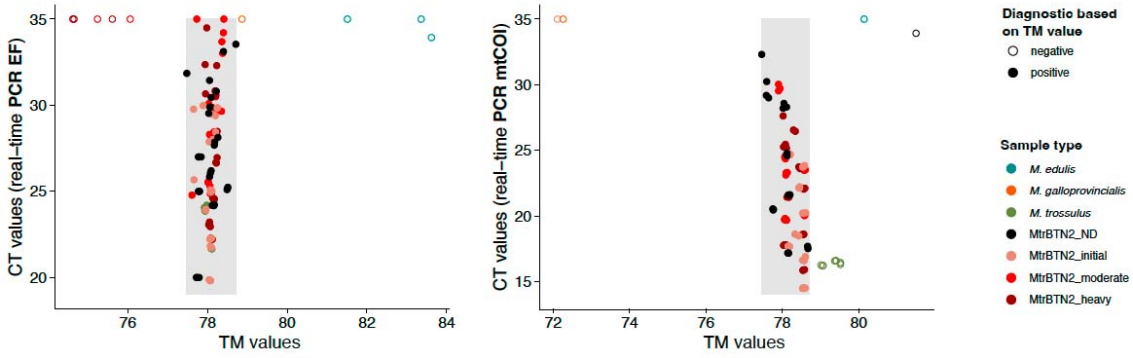
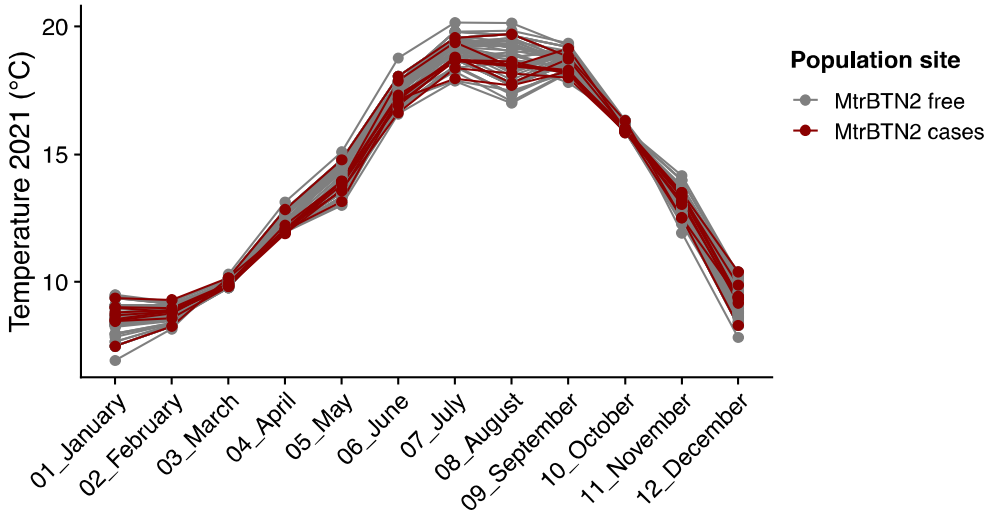


Figure S1: Diagram of screening and genotyping experimental steps.



668 **Figure S2: TM threshold determination for EF1α-i3 (left) and mtCOI-sub (right) real-**
669 **time PCR markers.** Grey rectangles correspond to TM threshold based on MtrBTN2
670 positive control samples (77.46<TM<78.71). Sample type colour and diagnostic based on
671 TM values are indicated in the legend. *M. edulis* and *M. galloprovincialis* correspond to
672 negative controls, *M. trossulus* are *M. trossulus* controls, and MtrBTN2 samples are positive
673 controls (ND: cancer stage not defined, initial: early stage, moderate: moderate stage, heavy:
674 late stage, see TableS2)
675
676
677



678

679 **Figure S3: Sampled sites temperature, recorded monthly in 2021.** Points correspond to a
680 temperature measurement, solid lines refer to MtrBTN2-free sites (grey) and MtrBTN2
681 affected sites (darkred).
682

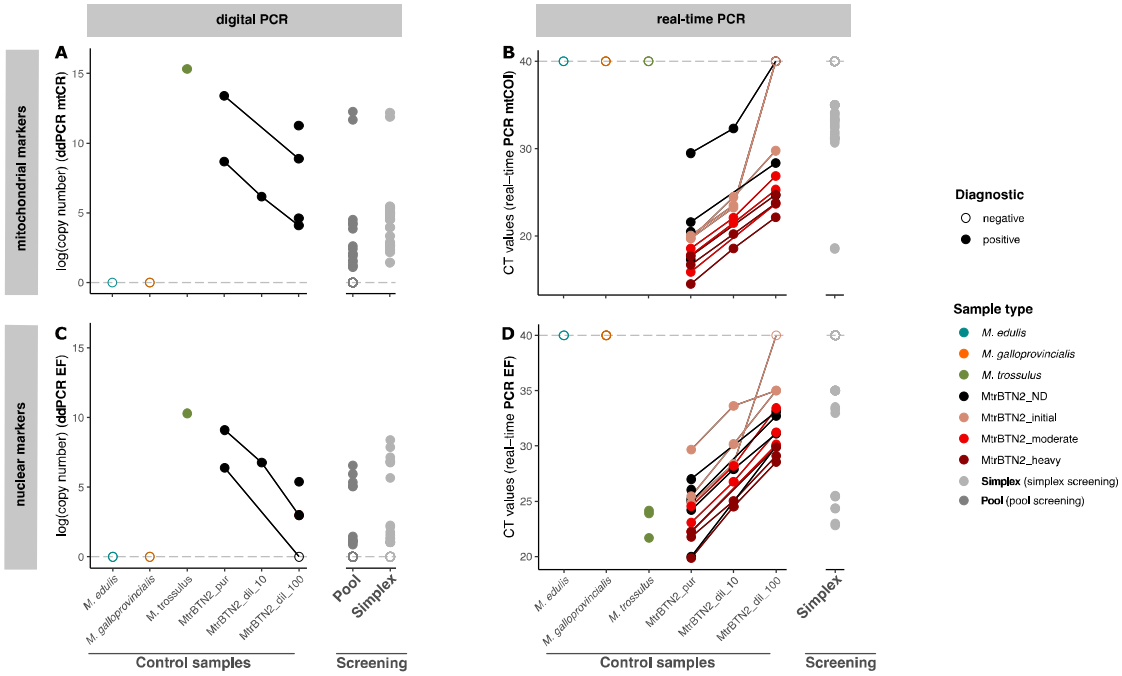


Figure S4: Real-time PCR and digital PCR results. (A) and (C) represent ddPCR results of EF and mtCR, and (B) and (D) represent real-time PCR results of EF1α-i3 and mtCOI-sub, respectively. Dotted lines correspond to the negative thresholds, 0 for no DNA copy detected by ddPCR and 40 for no amplification in real-time PCR. Solid lines in each graph refer to the same sample diluted to different concentrations (pure, x10, x100). *M. edulis* and *M. galloprovincialis* correspond to negative controls, *M. trossulus* are *M. trossulus* controls, and MtrBTN2 samples are positive controls (ND: cancer stage not defined, Initial: early stage, moderate: moderate stage, heavy: late stage, see TableS2). Screening results using both method and markers are shown for Pool and Simplex. Sample type colours and diagnostic status are indicated in the figure legend.

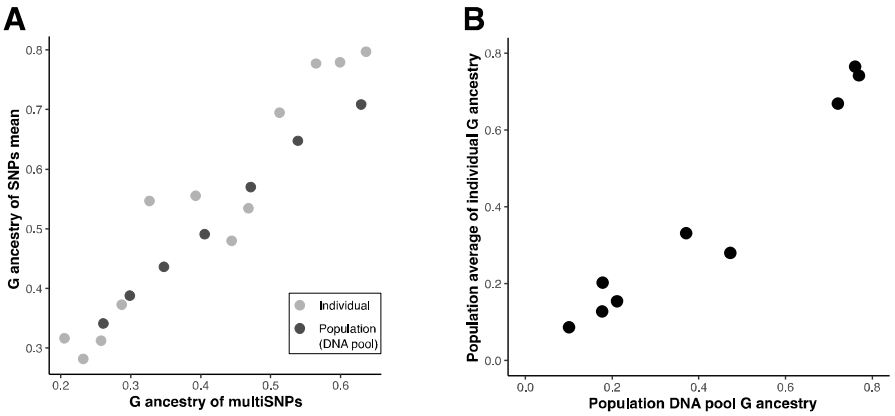


Figure S5: Genotyping method validation. (A) Positive correlation between multiSNPs G ancestry and the mean of the 10 single SNPs from a subset of Individual and Pool samples. **(B)** Positive correlation between population DNA pool G ancestry and the mean individual G ancestry of individuals from the same population.

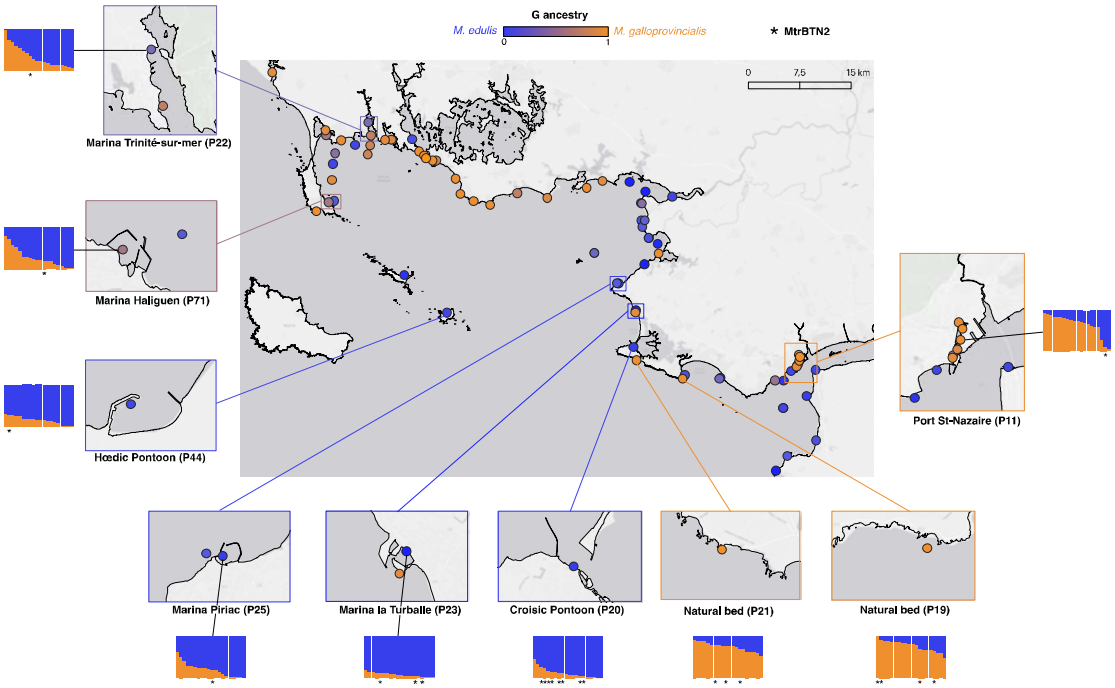


Figure S6: Population and individual ancestry composition across the sampled area. Zoomed-in sites correspond to sites with mussels affected by MtrBTN2. Barplots represent the estimated ancestry of individuals based on the 10SNPs multiplex tool. MtrBTN2 individuals are indicated by stars under each barplot. Coloured points represent the population average ancestry, as indicated in the figure's legend. Map origin: Ersi gray (light).

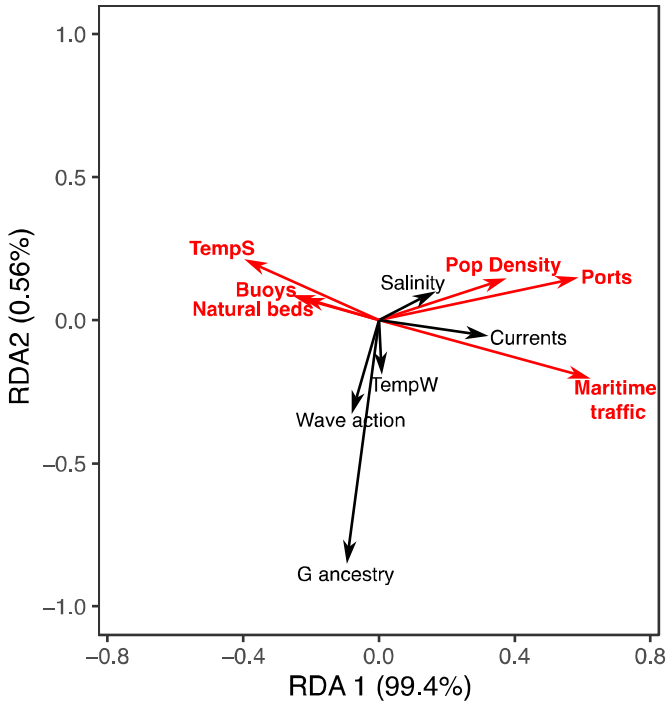


Figure S7: RDA analysis to assess overall relationship between presence of MtrBTN2 affected mussels and environmental factors. Arrows represent environmental variables. Red arrows correspond to significant variables obtained by forward selection with the ordiR2step() function (R software). Axis 1 correlates positively with the presence of MtrBTN2.

Supplementary Tables (➔ See excel TableS1-12)

Table S1: Primer and probe sequences used in real-time PCR and digital PCR.

Table S2: Control samples used to validate real-time PCR and ddPCR screening tools. All samples are from our laboratory sample collection. *M. edulis* and *M. galloprovincialis* correspond to negative controls, *M. trossulus* are *M. trossulus* controls, and MtrBTN2 samples are positive controls. MtrBTN2 stages were defined by cytology of the hemolymph, according to the following classification: early (<15% cancer cells), moderate (15-75% cancer cells), late (>75% cancer cells), ND when cancer stage was not defined.

Table S3: Bi-allelic nuclear SNPs used for the multiSNPs genotyping.

TableS4: Source and description of environmental variables.

TableS5: Description of sampled sites. Sources of environmental variables are provided in TableS4. Site: site name; MtrBTN2 prevalence: number of MtrBTN2 affected mussels; TempS: mean spring-summer temperature; TempW: mean autumn-winter temperature; temp: temperature; sal: salinity; vel: current; [fish-pass-other][19-21]_[1-4], fish: fishing boat density, pass: passenger boat density, other: recreational boat density, 19 to 21 correspond to the year of record, and 1 to 4 correspond to spring, summer, autumn, winter, respectively.

Table S6: Control sample results validate real-time PCR and ddPCR screening tools. EF_CopyNb: copy number for EF ddPCR marker; mtCR_CopyNb: copy number for mtCR ddPCR marker; CT: threshold cycle, TM: melting temperature for EF1α-i3 and mtCOI-sub real-time PCR markers. NP: not performed.

TableS7: Pooled and simplex screening results. Pool_CopyNb_EF: copy number for EF ddPCR marker (pooled screening); Pool_CopyNb_mtCR: copy number for mtCR ddPCR marker (pooled screening); CT: threshold cycle, TM: melting temperature for EF1α-i3 and mtCOI-sub real-time PCR markers; Simplex_CopyNb_EF: copy number for EF ddPCR marker (simplex screening); Simplex_CopyNb_mtCR: copy number for mtCR ddPCR marker (simplex screening). NP: not performed.

TableS8: Simplex screening results for positive samples. CT: threshold cycle, TM: melting temperature for EF1α-i3 and mtCOI-sub real-time PCR markers; PCR_result: marker name positive in real-time PCR; EF_CopyNb: copy number for EF ddPCR marker; mtCR_CopyNb: copy number for mtCR ddPCR marker; ddPCR_result: marker name positive in ddPCR; MtrBTN2: final diagnostic status. NP: not performed.

Table S9: Correspondence between pool and simplex screening. Pool_CopyNb_EF : copy number for EF ddPCR marker (pooled screening); Pool_CopyNb_mtCR : copy number for mtCR ddPCR marker (pooled screening); Pool_ddPCR_result: marker name positive in ddPCR (pool); Nb_ind_positive: number of individuals in the pool found positive for one or two marker in simplex screening; Simplex_PCR_result: marker name of the positive individual in real-time PCR; Simplex_ddPCR_result: marker name of the positive individual in ddPCR, negative if the sample did not amplify. NP: not performed.

Table S10: MultiSNPs genotyping results. G ancestry: fluorescence value; Ancestry: ancestry defined based on G ancestry values.

Table S11: Contingency table used for the Fisher exact test.

Table S12: Poisson regression models testing the effect of environmental variables on MtrBTN2 prevalence. Model results were obtained using glm() function in R. Variables explaining MtrBTN2 prevalence at the univariate level were selected for the multivariate model. TempS: spring-summer temperature.